

英語の適応的質問応答機能のための定型的解答文に対する正誤判定法の評価

A Method of Judging Stereotyped Answers for Questions and Answers and its Evaluation

國近 秀信*¹
Hidenobu KUNICHIKA

本田 実*^{2,†}
Minoru HONDA

平嶋 宗*³
Tsukasa HIRASHIMA

竹内 章*²
Akira TAKEUCHI

*¹ 九州工業大学大学院
Dept. of Creation Informatics
Kyushu Institute of Technology

*² 九州工業大学
Dept. of Artificial Intelligence
Kyushu Institute of Technology

*³ 広島大学大学院
Dept. of Information Engineering
Hiroshima University

This paper presents a method of judging answers to a question in the question and answer function in an intelligent CALL system. Both the questions and answers are English sentences. First, expected correct answers are generated by referring to the syntactic information in a question or a story. Then, the syntactic information of the answer is compared with that of the generated correct answers. We evaluated our method by using actual answer sentences written by learners.

1. はじめに

外国語学習において、テキストをリスニングまたは読解した後、その内容に関する質問に答えるという学習方法が一般的に利用されている。これまでに、質問応答をおこなう学習機能を備えた言語学習支援システム(たとえば、[Heift 01])が研究・市販されているが、あらかじめ用意された質問文を利用しているため、学習者の理解状態に合致した質問を出題することは大変困難である。

これに対して我々は、英語物語の表層的な内容を対象とし、動的に変化する学習者の理解状態に適した質問応答を実現することを目的に研究を行っている[國近 95]。この質問応答機能を実現するためには、解答文の正誤を判定するとともに、解答が誤っていた場合には、その誤りを同定する正誤判定機能が必要である。学習者に英文で解答させる場合は、学習者が必ずしも物語のテキスト文と同じ文を入力するとは限らないため、解答文の意味を抽出して正誤判断をおこなうことが必要である。しかし、学習者は、文形式の解答だけではなく、独特の形式を用いる場合があり、意味の比較だけでは正誤判定が困難な場合がある。本論文では、そのような定型的な解答文に対し、予想される正答を自動生成し、その正答と解答文とを比較することによる正誤判定法について述べる。以降の章では、まず質問応答機能の概要を述べ、次に正誤判定機能の実現法について説明する。その後、本正誤判定機能の評価について述べる。

2. 質問応答機能の概要

本学習機能は、中学校の教科書で使用されているレベルの英文を対象とし、物語の表層的な内容に関して質問応答をおこなう機能である。本学習機能では、システムが英語で質問をおこない、それに対して学習者が英語で解答をおこなう。対象となる学習者のレベルは、限定していない。学習者は「読む」または「聞く」などにより、その内容のある程度理解していることを前提とし、テキストを隠した状態で質問応答をおこなう。

本学習機能は、英文テキストから質問文を自動生成した後、各質問の難易度を算出し、質問応答履歴や学習者の理解状態を考慮して学習者に適切な質問文を選択・出題[國近 03]する。学習者の解答は、自然言語理解モジュールにより構文・意味情報が抽出され、構文的誤り同定および正誤判定がおこなわれる。

もし学習者が誤った入力をした場合は、学習者に熟考を促すという基本的姿勢で、ヒントや単純化した問題の提示をおこなう。

3. 正誤判定法

正誤判定では、原則として、解答に最低限必要な情報が含まれており、テキストの内容と矛盾しなければ正解と判断する。また、解答が、正解にも誤りにも判断できる場合は、正解の方を優先する。本章では、まず誤り情報とその利用について説明する。次に、想定される解答文の形式を整理する。その後、予想正答文との比較による正誤判定法について述べる。

3.1 誤り情報とその利用

誤り情報とは、誤りの種類、誤り箇所、誤った語句、および正しい語句のことを指す。学習者の解答文が誤っていた場合には、これらの情報を参照し、学習者に熟考を促すことを目的として段階的にヒントを提示する。つまり、最初は抽象的なヒントを与え、同じ誤りを繰り返す場合は、より具体的なヒントを提示する。たとえば、最も具体的なヒントとして、「"Tom"が"find"した対象は"the tennis court"ではなく"the school"です」を学習者に与える。

3.2 解答文の形式

本質問応答機能の質問文の形式は、一般疑問文と特殊疑問文であり、それぞれ「Yes/No」と「疑問詞」に対応するテキストの内容が「解答に最低限必要な情報」である。本研究ではこれを必須情報と呼ぶ。我々は、中学校の英語の教科書や参考書を調査し、解答文として入力される可能性のある解答文の形式を整理した。以下、本質問応答機能で取り扱う解答形式を示す。

- (f1) 必須情報のみ
- (f2) Yes/No, 代名詞 + {do(does, did), 助動詞, be 動詞}
- (f3) 必須情報 + {do(does, did), 助動詞, be 動詞}
- (f4) 解答指示語 + be 動詞 + 必須情報
- (f5) 平叙文
- (f6) Yes/No, 平叙文

3.3 実現法

本正誤判定では、まず解答形式の検査をおこない、その後、解答文の内容の確認をおこなう。

(1) 解答形式の検査

自然な質問応答が成立するためには、質問・解答間の適切な対応関係がとれていることが必要である。本研究では、教科

連絡先: 國近秀信, kunitika@ci.kyutech.ac.jp

†現在, セイコーエプソン株式会社

書や文法書等を参考にして、質問形式と解答形式の対応関係を調査・整理した結果を利用し、一般疑問文に対する解答文が3.2節の(f2)および(f6)の形式になっているかどうか、また特殊疑問文に対する解答文が(f3)-(f5)の形式になっているかどうかを検査する。

(2) 正答の自動生成と正誤判定

本研究では、3.2節で述べた解答文の形式のうち、定型的な解答文として扱うことができる f1 から f4 までを対象とし、予想正答文と学習者の解答文の単語および構文情報を比較することで、正誤判定および誤りの検出をおこなう。その際には、必須情報以外については、仕組みの単純化のために、単純に語句の比較をおこなう。また、必須情報については、修飾語句の省略を許すために、主要部(head)の原形同士の比較をおこなう。もし、必須情報が修飾語句以外であり、修飾語句を省略しない場合は、「質問生成元と同一の修飾語句に限る」という方針で検査をおこなう。ここで、必須情報の修飾語句の省略を許可する理由は、質問応答の解答文の中で修飾語句が欠けていたとしても一概に誤りとは言いきれないためである。

以下に、各解答形式に対する誤り検出の方法を説明する。

(jf1) 必須情報のみ

解答が単語のみのとき、それが必須情報と一致するかどうかを主要部の原形同士の比較により判断する。

(jf2) Yes/No, 代名詞 + {do(does, did), 助動詞, be 動詞}

質問文中から主語と助動詞を取り出して、主語を代名詞に置換することにより正答の雛型を生成し、解答文と比較する。ただし、質問文と同一の助動詞が使用できない場合もあるため、変換知識を用意している。

(jf3) 必須情報 + {do(does, did), 助動詞, be 動詞}

質問文中から助動詞を取り出して、必須情報の後に付加することで正答の雛型を生成し、解答文との比較をおこなう。

(jf4) 解答指示語 + be 動詞 + 必須情報

たとえば、「Who is Jean's brother?」という質問に対し、「It is Ted.」と解答する場合は考えられる。このような形式の場合、必須情報に対応する解答指示語(代名詞)と be 動詞とを選択することにより正答の雛型を生成し、解答文と比較する。

4. 評価

我々は、本研究の正誤判定法が実際の解答文に適用可能かどうかを確認するために、調査をおこなった。

4.1 調査方法

中学校の英語の教科書にある物語から 6 文を取り出し、質問文自動生成機能を用いて質問文を生成した。その中から、一般疑問文および特殊疑問文を各 5 問選択した。一般疑問文のうち、2 問が No と答えるべき質問である。調査の際には、まず 2 分間物語を読んで内容を理解してもらい、その後物語を見ずに各問題に解答してもらった。解答文の形式および数については自由とした。被験者は、本学の大学生および大学院生 32 名である。解答文の調査の後に、本正誤判定法で対象としている解答形式(f1)から(f4)の解答文について、正しく正答および誤答を判断できるかどうかを確認した。

4.2 結果と考察

調査の結果、一般疑問文については 165 の解答文が得られ、特殊疑問文については 163 の解答文が得られた。解答文の形式毎の数を調べたところ、本手法が対象とする解答形式(f1)から(f4)の数は、一般疑問文では 130 文、また特殊疑問文では 75

表 1 解答文の数

	一般疑問文	特殊疑問文
対象とする解答形式 (1問あたりの数)	130 (26)	75 (15)
正しく判定した解答文	121	73
正しく判定した割合 (%)	93	97

表 2 解答文のバリエーション数

	一般疑問文	特殊疑問文
対象とする解答形式 (1問あたりの数)	22 (4.4)	54 (10.8)
正しく判定した解答文	20	52
正しく判定した割合 (%)	91	96

文であった。このことより、本手法は、定型的解答文が多い一般疑問文だけでなく、平叙文で解答することが多い特殊疑問文に対しても半数弱の 46% (163 文中 75 文) の解答文を対象とできることがわかる。

表 1 に、本正誤判定法で対象としている解答形式の解答文と正しく判定された解答文の数を示す。これらには、同じ解答も含まれている。異なる解答を数えた結果を、表 2 に示す。表 2 より、(f1)から(f4)の解答形式は、解答のバリエーションが比較的少ないとはいえ、まったく同じ文ばかりではなく、ある程度の数のバリエーションが存在することがわかる。また、表 1 および表 2 より、本手法は、解答形式(f1)から(f4)について、実際の質問応答の際に入力されるほとんどの解答文を正しく正誤判定できるということができる。

正誤判定が失敗した原因は、「同義語に対応できない」、および「必須情報については単数/複数の属性が異なっても比較は成功するが、その違いによる他の語の変形に対応できない」であった。

5. おわりに

本論文では、質問応答機能のための予想正答文の自動生成、それを利用して正誤判定および誤り同定をおこなう機能の実現、およびその評価について述べた。実験の結果、本手法を利用することで、一般疑問文の解答文に対しては 79%、また特殊疑問文の解答文に対しては 46% をカバーし、それらのうちのほとんどを正しく正誤判定できることが確認できた。

今後は、より多くの解答文の正誤判定ができるようにするため、意味情報の比較により正誤判定をおこなう手法を実現することが必要である。また、本論文で述べた評価は、大学生および大学院生が書いた英文を用いているため、他のレベルの学習者の英文を用いて評価をおこなうことも必要である。

参考文献

- [Heift 01] Heift, T. & Nicholson, D.: "Web Delivery of Adaptive and Interactive Language Tutoring", International Journal of AIED, Vol.12, pp.310-324, 2001.
- [國近 95] 國近秀信, 竹内章, 大槻説平: 知的学習支援システム HELEN のオーサリング環境, 教育システム情報学会誌, Vol.12, No.1, pp.52-62, 1995.
- [國近 03] 國近秀信, 宇留島稔, 平嶋宗, 竹内章: "英語長文読解学習のための適応的出題制御", 教育システム情報学会研究報告, Vol.18, No.3, pp.7-12, 2003.